UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/660,635	09/13/2000	Alan Lawrence Higgins	D-ACD-W019 HIGGINS	7304

28581 7590 07/12/2004

DUANE MORRIS LLP
100 COLLEGE ROAD WEST, SUITE 100
PRINCETON, NJ 08540-6604

EXAMINER	
WOZNIAK, JAMES S	

ART UNIT	PAPER NUMBER
2655	

DATE MAILED: 07/12/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/660,635	HIGGINS ET AL.	
	Examiner	Art Unit	
	James S. Wozniak	2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *6/1/2004*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-17* is/are pending in the application.

4a) Of the above claim(s) *1-5* is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *6-17* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☒ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *13 September 2000* is/are: a)☐ accepted or b)☒ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

Detailed Action

Response to Amendment

1. In response to the office action from 2/26/2004, the applicant has submitted an amendment, filed 6/1/2004, affirming the provisional election of claims 6-17 for prosecution on the merits, while arguing to traverse the art rejection based on the limitation regarding a fixed-length feature vector that is independent of word order or speaking rate *(Amendment, Page 2)*, challenging the motivation of reference combination in the rejection of Claim 6 *(Amendment, Page 2)*, and requesting references in support of official notice taken in the office action (*Amendment, Page 3)*. Applicant's arguments have been fully considered, however the previous rejection is maintained due to the reasons listed below in the response to arguments and altered only in regards to the applicant's request for documentation in support of the official notice taken with respect to Claim 10.

Response to Arguments

2. Applicant's arguments have been fully considered but they are not persuasive for the following reasons:

- With respect to the objection of **Claims 7 and 14**, the applicant argues that the Baum-Welch algorithm can be alternatively referred to as "Baum-Welsh" *(Amendment, Page 2)*. However, as is well known in the art, Leonard Baum and

Lloyd *Welch* developed this algorithm and the title represents both developer names. Therefore, referring to the algorithm as "Baum-Welsh" would be incorrect and the objection of Claims 7 and 14 is maintained.

- With respect to **Claim 6**, the applicant argues that Kuhn et al does not teach or suggest that a fixed-length vector that is independent of the order of words spoken or the speaking rate and states that no motivation for combining Kuhn et al *(U.S. Patent: 6,343,267)* with Vysotsky et al *(U.S. Patent: 5,832,063)* has been supplied *(Amendment, Page 2)*. In regards to the argument regarding a fixed-length feature vector that is independent of the order of words spoken or the speaking rate, Kuhn teaches that the supervector upon which dimensionality reduction is performed contains an Eigenspace that includes all of the training data for a particular speaker *(Col. 6, Line 62- Col. 7, Line 35)*. Thus, since all possible utterances of a user are included within this space, the order of words spoken by a speaker in a speaker recognition process would not be important.

  Also as noted by Kuhn, by using a maximum likelihood technique, a supervector is selected that is most consistent with input speech (Col. 9, Lines 4-13), so that regardless of the way a user speaks (including speed) an appropriate supervector would be selected for speaker and speech recognition that would further undergo the aforementioned dimensionality reduction (Col. 6, Lines 36-45).

  Since Kuhn has made no mention of variable length dimensionally reduced supervectors, it is presumed that the supervectors are of a fixed length.

In response to applicant's argument that there is no suggestion to combine the references, the examiner recognizes that obviousness can only be established by combining or modifying the teachings of the prior art to produce the claimed invention where there is some teaching, suggestion, or motivation to do so found either in the references themselves or *in the knowledge generally available to one of ordinary skill in the art.* See *In re Fine*, 837 F.2d 1071, 5 USPQ2d 1596 (Fed. Cir. 1988) and *In re Jones*, 958 F.2d 347, 21 USPQ2d 1941 (Fed. Cir. 1992). In this case, it is clearly stated in the first office action that the reason for combining Vysotsky with Kuhn is: "to provide for adaptive speaker recognition if speech rate is altered especially in the enrollment process since all speakers do not speak at the same rate of speed, or if the order of words has been altered, for example, if the order of numbers within a voice password is changed" *(FAOM, Page 5)*. "To obtain the invention as specified in Claim 6" is not meant as a reason for motivation, it is merely a summation of the preceding motivational statement. Furthermore, Kuhn additionally provides a reason for combination noting that the disclosed dimensionality reduction allows for "considerable flexibility and computational economy" *(Col. 3, Lines 41-52, and compression within systems having limited memory and processor resources, Col. 7, Lines 31-35)* in a speaker verification system.

Thus, the rejection of **Claim 6** is maintained.

- With respect to **Claim 10**, the applicant requests a reference in support of official notice. With respect to the bridging word, "ti", Gandhi et al *(U.S. Patent:*

5,687,287) teaches the use of the word "ty" in a user password *(Col. 8, Lines 7-27)*, which is a functional equivalent of "ti". Also, Gandhi teaches the concatenation of word models to form a feature vector to be used for speaker enrollment and verification *(Col. 4, Line 53- Col. 5, Line 35)*. Thus, since Gandhi teaches the limitations for which official notice was taken, the rejection of Claim 10 is maintained.

- **Dependent claims, 7-9 and 11-17,** have not been argued with respect to the merits in regards to the art rejection and are dependent upon rejected independent claims, thus the rejection of these claims is also maintained.

Election/Restrictions

3. Thus, **Claims 1-5** are withdrawn from further consideration pursuant to 37 CFR 1.142(b) as being drawn to a nonelected Invention 1.

Specification

4. The disclosure is objected to because of the following informalities: "Baum-Welsch" should be corrected to read --Baum-Welch--, for example on Page 7, Line 21.

Appropriate correction is required.

Drawings

5. The drawings are objected to because "Baum-Welsch" should be corrected to read --Baum-Welch--, for example Fig. 1, Element 11.

The objection to the drawings will not be held in abeyance.

Claim Objections

6. Claims 7 and 14 are objected to because of the following informalities: "Baum-Welsh" should be corrected to read --Baum-Welch--. Appropriate correction is required.

Claim Rejections - 35 USC § 103

7. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

8. **Claims 6-9** are rejected under 35 U.S.C. 103(a) as being unpatentable over Vysotsky et al *(U.S. Patent: 5,832,063)* in view of Kuhn et al *(U.S. Patent: 6,343,267)*.

With respect to **Claim 6,** Vysotsky discloses:

In a method of automatically verifying a speaker as matching a claimed identity wherein enrollment speech data of a known speaker is compared with test data, including the steps of

processing spoken input enrollment speech data and test speech data into speech signals into series of frames of digital data representing the input speech, analyzing the speech frames by a speaker verification module which compares the enrollment and test features and generates respective match scores therefrom, and determining whether the test speech corresponds with the enrollment speech based upon the match scores, the improvement wherein:

The step of processing the spoken input enrollment and test speech data includes performing a feature extraction process on the enrollment and test speech data *(feature extraction, Col. 7, Lines 38-45)*; and

The step of analyzing the speech frames by comparison includes computing a weighted Euclidean distance between the feature vectors by a discriminative analysis *(Euclidean distance used in speech recognition, Col. 8, Lines 44-47, and discriminative analysis of feature vectors, Col. 11, Lines 58-63)*.

Vysotsky does not specifically teach the ability to convert variable input to fixed-length feature vectors that are independent of the order of words spoken or the speaking rate, however Kuhn discloses:

Ability to convert variable input to fixed-length feature vectors *(dimensionality reduction, Col. 6, Lines 62-64, and Col. 7, Lines 23-26)* that are independent of the order of words spoken or the speaking rate *(adaptive speaker models in the form of a supervector that is fully populated with parameter values for recognizing speech, thus word order would not be important since all parameter values would be contained within the Eigenspace, Col. 9, Lines 41-51. Also, using a maximum likelihood technique, a supervector is selected which is most consistent with input speech, Col. 9, Lines 4-13, so that, regardless of the rate of speech, a*

proper supervector would be selected for speaker and speech recognition. Furthermore, using singular value decomposition the supervector dimensionality is reduced, Col. 6, Lines 36-45).

Vysotsky and Kuhn are analogous art because they are from a similar field of endeavor in speech and speaker recognition. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to combine the use of a supervector containing all speech parameter values within an Eigenspace for speaker adaptation as taught by Kuhn with the speaker verification method through feature extraction and the computation of Euclidean distances between feature vectors as taught by Vysotsky to provide for adaptive speaker recognition if speech rate is altered, especially in the enrollment process since all speakers do not speak at the same rate of speed, or if the order of words has been altered, for example, if the order of numbers within a voice password is changed . Therefore, it would have been obvious to combine Kuhn with Vysotsky for the benefit of obtaining a method of adaptive speaker recognition capable of detecting speech regardless of speech rate or word order, to obtain the invention as specified in Claim 6.

With respect to **Claim 7**, Vysotsky teaches speaker verification through feature extraction and computing Euclidean distances between feature vectors as applied to Claim 6, which also utilizes a method of HMM adaptation through a Gaussian estimation *(Col. 8, Lines 40-43).* Vysotsky does not specifically further suggest that the aforementioned method of Gaussian estimation utilizing a Baum-Welch algorithm, however it would have been obvious to one of ordinary skill in the art, at the time of invention, to specifically utilize a Baum-Welch algorithm for HMM parameter adaptation since it is a well-known and common means of HMM parameter estimation in the art of speech recognition and has readily available software.

With respect to **Claim 8**, Vysotsky further recites:

The predetermined number of vocabulary words comprises five words, namely, "four", "six", "seven", "nine", and "ti" *(voice password for user verification as a string of digits comprising a word, Col. 11, Lines 45-55)*.

It also would have been obvious to one of ordinary skill in the art, at the time of invention that a model set (vocabulary) relating to a digit string voice password would contain four, six seven, nine, and in the case of a password such as "470" or "four seventy", ti, since a password-based speaker verification system would commonly utilize number-related vocabulary words in order to recognize numbers in a password sequence.

With respect to **Claim 9**, Vysotsky teaches the method of speaker verification through feature extraction and computing Euclidean distances between feature vectors, which also utilizes a method of HMM adaptation through a Gaussian estimation and contains a vocabulary corresponding to digits in a numerical password as applied to Claim 8. Vysotsky does not specifically further suggest a feature vector as a concatenation of state mean vectors as recited in Claim 9, however it would have been obvious to one of ordinary skill in the art, at the time of invention, to concatenate the mean vectors of the adapted HMMs in order to create a feature vector related to an entire password sequence for speaker verification which can provide instant recognition of an entire password at once, instead of recognition of individual password digits in sequence.

9. **Claims 10-17** are rejected under 35 U.S.C. 103(a) as being unpatentable over Vysotsky et al in view of Gandhi et al *(U.S. Patent: 5,687,287)*.

With respect to **Claim 10**, Vysotsky discloses:

In a voice verification system for dividing speech utterances into speech frames and analyzing the frames independently to verify one speaker's voice as compared to another's, the improvement therewith of a method for verifying a speaker's voice by subjecting the speaker to an enrollment test for verification based upon the premise that speech utterances are a fixed set of words arranged in a randomized order, comprising the steps of:

Causing said speaker to enroll by uttering from a vocabulary a predetermined number of combined words each word indicative of a number between one to nine and at least one bridging word "ti" *(user enrollment of a voice password comprising a string of digits, Col. 11, Lines 45-55. Also, it would have been obvious to one of ordinary skill in the art, at the time of invention that a model set (vocabulary) relating to a voice password would contain four, six seven, nine, and in the case of a password such as "470" or "four seventy", ti, since a password-based speaker verification system would commonly utilize number-related vocabulary words in order to recognize numbers in a password sequence, as is evidenced by Gandhi (*use of the word "ty" in a user password, *Col. 8, Lines 7-27*, which is a functional equivalent of "ti *))*.

Adapting the parameters of a set of word models for said vocabulary words based upon input speech data to provide adapted word models *(creating speaker dependent word models, Col. 8, Lines 40-43)*.

Vysotsky does not specifically suggest a feature vector as a concatenation of state mean vectors, however it would have been obvious to one of ordinary skill in the art, at the time of invention, to concatenate the mean vectors of the adapted HMMs, as is well known in the art, in order to create a feature vector related to an entire password sequence for speaker verification

which can provide instant recognition of an entire password at once, instead of recognition of individual password digits in sequence, as is evidenced by Gandhi *(concatenation of word models to form a feature vector to be used for speaker enrollment and verification , Col. 4, Line 53- Col. 5, Line 35)*. Furthermore, Vysotsky and Gandhi are directed towards a similar field of endeavor in speaker recognition, and would have been obvious for combination in order to obtain the capability of recognizing an entire spoken password sequence as is noted above, to obtain the invention as specified in Claim 10.

With respect to **Claim 11**, Vysotsky further discloses:

Comparing said feature vector obtained from said enrollment with a feature vector obtained from a speech test to determine the identity of said one speaker voice *(voice verification through comparison of feature vectors corresponding to a voice password to identify either a true or impostor speaker, Col. 11, Lines 45-63)*.

With respect to **Claim 12**, Vysotsky further recites:

Feature comparison is implemented by subjecting said vectors to a weighted Euclidean Distance computation *(Euclidean distance used in speech recognition, Col. 8, Lines 44-47)*.

With respect to **Claim 13**, Vysotsky further discloses:

The words are indicative of numbers, namely, "four", "six", "seven", "nine", and "ti" *(voice password for user verification as a string of digits comprising a word, Col. 11, Lines 45-55)*.

It also would have been obvious to one of ordinary skill in the art, at the time of invention that a model set (vocabulary) relating to a digit string voice password would contain four, six seven, nine, and in the case of a password such as "470" or "four seventy", ti, since a password-

based speaker verification system would commonly utilize number-related vocabulary words in order to recognize numbers in a password sequence.

With respect to **Claim 14**, Vysotsky discloses the speaker verification system featuring speaker enrollment through a voice password and adaptive word models as applied to Claim 10. Vysotsky does not specifically suggest model adaptation implementing a Baum-Welch algorithm, however it would have been obvious to one of ordinary skill in the art, at the time of invention, to specifically utilize a Baum-Welch algorithm for HMM parameter adaptation since it is a well-known and common means of HMM parameter estimation in the art of speech recognition and has readily available software.

With respect to **Claim 15**, Vysotsky further discloses a feature vector matrix used for comparison to input speech feature vectors for voice identification *(Col. 8, Lines 4-7)*. Also, it would have been obvious to one of ordinary skill in the art, at the time of invention that the dimensionality of this matrix could have a value of 1568, for instance, in a 49X32 or other such matrix configuration, based on desired system settings.

With respect to **Claim 16**, Vysotsky further discloses:

Forming said feature vector for each speaker using the difference in vectors between a first and second speaker channel *(speaker reference model adapted and thus formed according to changes in speaker and channel coupling, Col. 11, Line 64- Col. 12, Line 1)*.

With respect to **Claim 17**, Vysotsky teaches the speaker verification system featuring speaker enrollment through a voice password and adaptive word models responsive to changes between speaker channels as applied to Claim 16. Vysotsky does not specifically suggest the approximation of speech with white noise channel differences in deriving speaker features as

recited in Claim 17, however it would have been obvious to one of ordinary skill in the art, at the time of invention, to include white noise approximation in speech features, since white noise is common to telephone communication channels, and thus, should be included within the speaker feature vectors modeled from speech inputs to the communication channels to better approximate their expected characteristics.

Conclusion

10. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure:

- Sukkar *(U.S. Patent: 5,613,037)*- teaches a spoken digit recognition system that utilizes the concatenation of variable length utterances to form fixed-length feature vectors for recognition.

11. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event,

however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

12. Any inquiry concerning this communication or earlier communications from the examiner should be directed to James S. Wozniak whose telephone number is (703) 305-8669 and email is James.Wozniak@uspto.gov. The examiner can normally be reached on Mondays-Fridays, 8:30-4:30.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Doris To can be reached at (703) 305-4827. The fax/phone number for the Technology Center 2600 where this application is assigned is (703) 872-9306.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the technology center receptionist whose telephone number is (703) 306-0377.

W. R. YOUNG
PRIMARY EXAMINER

James S. Wozniak
6/30/2004